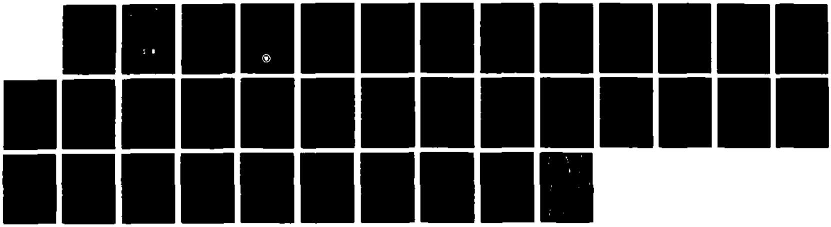
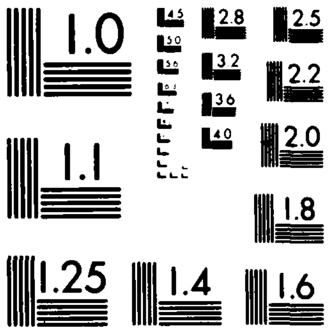
AD-A186 026 ON SIMULTANEOUS ESTIMATION OF THE NUMBER OF SIGNALS AND FREQUENCIES UNDER..(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS Z D BAI ET AL. DEC 86 TR-86-37 1/1

UNCLASSIFIED AFOSR-TR-87-1016 F49620-85-C-0008 F/G 12/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

AD-A186 026

Unclassified

DTIC FILE COPY

1

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1016	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On simultaneous estimation of the number of signals and frequencies under a model with multiple sinusoids		5. TYPE OF REPORT & PERIOD COVERED ~~Technical~~ Journal - December 1986
		6. PERFORMING ORG. REPORT NUMBER 86-37
7. AUTHOR(s) Z.D. Bai, P.R. Krishnaiah and L.C. Zhao		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008 N00014-83-K-0387
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis Fifth Floor Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Office of Naval Research NM		12. REPORT DATE December 1986
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as 11		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE OCT 1 6 1987 S D D

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

consistency, estimation, frequencies, multiple sinusoids, signals

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, the authors considered the problem of estimation of the frequencies and the number of signals under a signal processing model with multiple sinusoids. The frequencies are estimated with eigenvariation linear prediction method. The number of signals is estimated with an in-

DD FORM 1 JAN 73 1473

20

formation theoretic criterion. The strong consistency of the estimates of the frequencies and the number of signals is also established. Also, a modification of forward backward linear prediction method is suggested to yield consistent estimators of the frequencies.

AFOSR-TR- 87-1016

ON SIMULTANEOUS ESTIMATION OF THE NUMBER OF SIGNALS AND FREQUENCIES UNDER A MODEL WITH MULTIPLE SINUSOIDS*

Z.D. Bai, P.R. Krishnaiah and L.C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

ON SIMULTANEOUS ESTIMATION OF THE NUMBER OF SIGNALS AND FREQUENCIES UNDER A MODEL WITH MULTIPLE SINUSOIDS*

Z.D. Bai, P.R. Krishnaiah and L.C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

December 1986

Technical Report No. 86-37

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* Research sponsored by the Office of Naval Research, under Contract N00014-83-K-0387 and the Air Force Office of Scientific Research (AFOSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

ON SIMULTANEOUS ESTIMATION OF THE NUMBER OF SIGNALS AND FREQUENCIES UNDER A MODEL WITH MULTIPLE SINUSOIDS*

ABSTRACT

In this paper, the authors considered the problem of estimation of the frequencies and the number of signals under a signal processing model with multiple sinusoids. The frequencies are estimated with eigenvariation linear prediction method. The number of signals is estimated with an information theoretic criterion. The strong consistency of the estimates of the frequencies and the number of signals is also established. Also, a modification of forward backward linear prediction method is suggested to yield consistent estimators of the frequencies.

Key words and phrases: consistency, estimation, frequencies, multiple sinusoids, signals.

* Research sponsored by the Office of Naval Research, under Contract N00014-83-K-0387 and the Air Force Office of Scientific Research (AFOSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

1. INTRODUCTION

Consider the model

$$y(n) = \sum_{k=1}^{t_0} a_k \exp(i\omega_k n) + w(n), \quad n = 1, 2, \ldots, N \tag{1.1}$$

where $a_k \neq 0$ are unknown amplitudes, $\omega_k \in (0,2\pi)$, $(k = 1, 2, \ldots, t_0)$, are unknown frequencies and $i = \sqrt{-1}$. Usually we assume that the noise $w(n)$ has mean zero and finite variance σ^2. The above model is of interest in tne area of signal processing. Under the above model, it is of interest to estimate the unknown parameters. Even when t_0 is known, it is difficult to find the least square estimates of a_k's and ω_k's since it would involve solving a system of nonlinear equations with exponential functions. To avoid this difficulty, various methods have been developed in the literature, such as linear prediction (LP) method, those based upon using principal eigenvectors of estimated cross-spectral density matrices (Liggett (1973)) and the forward-backward linear prediction (FBLP) method (Nuttall (1976) and Ulrych and Clayton (1976)).

In the LP procedure, we have still to solve a polynomial equation whose degree would be rather high although we do not solve a system of exponential equations. Also, it does not work well for the case of low signal-to-noise ratio (SNR). Tufts and Kumaresan (1982) made modification to the original FBLP method, and showed by simulation that the modified FBLP works much better than the original FBLP when the SNR is relatively low. However, it still involves solving a high degree polynomial equation.

In the present paper, we investigate the estimation of both the number of signals and the amplitudes and frequencies of the signals, and study a method which we refer to as equivariation linear prediction

(EVLP). Using this approach, we can at the same time find the estimates of number of signals and the frequencies and then using the usual least square method to get the estimates of the amplitudes. In this method, we need only to solve a polynomial equation with the lowest degree. In Section 2 we shall state this method. In Section 3 we shall prove the strong consistency of this procedure. In Section 4 we give the limiting distribution of various estimates, given in previous sections. In Section 5 we qive a further discussion on the FBLP and the modified one. In this section we will show the reason why (theoretically) the modified FBLP works better than the original one when the SNR is relatively low and the sample size is small, and show that these two methods will become equivalent when the sample size goes to infinity. More importantly, we shall point out that both these two methods do not provide consistant estimation of the frequencies, and we shall pose a further modification on FBLP such that the procedure is consistent.

2. DETERMINATION OF THE NUMBER OF THE FREQUENCIES AND ESTIMATION OF THE FREQUENCY PARAMETERS

Suppose that the data sequence $y(n)$ is given by the formulas

$$y(n) = \sum_{j=1}^{t_0} a_j \exp(i\omega_j n) + w(n), \quad n = 1, 2, \ldots, N, \tag{2.1}$$

where $i = \sqrt{-1}$, $\{a_j\}$ is a set of unknown complex amplitudes, $\{\omega_j\}$ is a set of unknown angular frequencies, and $\{w(n)\}$ is a sequence of i.i.d. complex random noise variables such that

$$Ew(1) = 0, \quad Ew(1)\overline{w(1)} = \sigma^2, \quad E|w(1)|^4 < \infty \tag{2.2}$$

with σ^2 unknown. We assume that $\omega_j \neq \omega_k$ if $j \neq k$, and $\omega_j \in (0,2\pi)$ for all j.

In this paper, we are primarily interested in determining t_0 and estimating the frequency parameters ω_j. Once ω_j's are accurately determined, the a_j's can be found by a linear least squares fit to the data.

To determine t_0, the true number of different angular frequencies, it is prescribed a prior that $t_0 \leq T < \infty$.

For any nonnegative integer $t \leq T$, write complex vector $\underset{\sim}{b}^{(t)}$ as

$$\underset{\sim}{b}^{(t)} = (b_0^{(t)}, \ldots, b_t^{(t)})'. \tag{2.3}$$

Put*

$$S_t = \min\left\{\frac{1}{N-t} \sum_{n=t+1}^{N} \left| \sum_{\ell=0}^{t} b_\ell^{(t)} \overline{y(n-\ell)} \right|^2 : \|\underset{\sim}{b}^{(t)}\| = 1\right\},$$

$$t = 0, 1, 2, \ldots, T, \tag{2.4}$$

---

* Rao (1986) also remarked about finding S_t, for given t, to estimate the frequencies.

where $\|\underset{\sim}{b}^{(t)}\| = (\sum_{\ell=0}^{t} |b_\ell^{(t)}|^2)^{1/2}$. Take C_N satisfying the following conditions:

$$\lim_{N\to\infty} C_N = 0 \quad \text{and} \quad \lim_{N\to\infty} \sqrt{N}\, C_N/\sqrt{\log\log N} = \infty. \tag{2.5}$$

Then we can find a nonnegative integer $\hat{t}_0 = \hat{t}_N \leq T$ which minimize

$$R_t = S_t + tC_N, \quad t = 0, 1, \ldots, T, \tag{2.6}$$

and use $\hat{t}_0$ as an estimate of t_0.

Further, we can find a unit $(\hat{t}_0 + 1)\text{x}1$ complex vector $\hat{\underset{\sim}{b}} = (\hat{b}_0, \ldots, \hat{b}_{\hat{t}_0})'$

$$S_{\hat{t}_0} = \frac{1}{N-\hat{t}_0} \sum_{N=\hat{t}_0+1}^{N} |\sum_{\ell=0}^{\hat{t}_0} \hat{b}_\ell \overline{y(n-\ell)}|^2. \tag{2.7}$$

Let $\hat{\rho}_j \exp(i\hat{\omega}_j)$, $j = 1, \ldots, \hat{t}_0$, be the $\hat{t}_0$ solutions of

$$\hat{B}(z) \triangleq \sum_{j=0}^{\hat{t}_0} \hat{b}_j z^j = 0, \tag{2.8}$$

where $\hat{\rho}_j > 0$, $\hat{\omega}_j \in [0, 2\pi)$, $j = 1, \ldots, \hat{t}_0$. Then we use $\hat{\omega}_j$'s as estimates of ω_j's.

Note that if $\hat{\underset{\sim}{b}}^{(t)}$ satisfies

$$S_t = \frac{1}{N-t} \sum_{n=t+1}^{N} |\sum_{\ell=0}^{t} \hat{b}_\ell^{(t)} \overline{y(n-\ell)}|^2, \quad t = 0, 1, \ldots, T, \tag{2.9}$$

then S_t is the smallest eigenvalue of the matrix

$$\hat{\Gamma}^{(t)} = (\hat{\gamma}_{\ell m}^{(t)}), \quad \ell, m = 0, 1, \ldots, t, \tag{2.10}$$

and $\hat{\underset{\sim}{b}}^{(t)}$ is the corresponding unit eigenvector, where

$$\hat{\gamma}_{\ell m}^{(t)} = \frac{1}{N-t} \sum_{n=t+1}^{N} y(n-\ell)\overline{y(n-m)}, \qquad \ell, m = 0, 1, \ldots, t. \tag{2.11}$$

The above method combines the procedure of parameter estimation with the detection procedure of the parameter number. As shown in Section 3, $(\hat{t}_0, \{\hat{\omega}_j, j \leq \hat{t}_0\})$ is a strong consistent estimate of $(t_0, \{\omega_j, j \leq t_0\})$ under the condition (2.2). Besides, as an estimate of σ^2, $S_{\hat{t}_0}$ is also strongly consistent. In Section 4, we obtain the limiting distributions of $S_{\hat{t}_0}$ and $\{\hat{\omega}_j, j \leq \hat{t}_0\}$.

3. STRONG CONSISTENCY OF THE DETECTION AND ESTIMATION PROCEDURES

In this section, we establish the following.

THEOREM 3.1. Suppose that $\{w(n)\}$ is an i.i.d. sequence such that (2.2) holds. Then with probability one for N large, the following results hold:

(i) $\hat{t}_N = t_0$,

(ii) there exist a unique $(t_0 + 1) \times 1$ unit vector $\hat{\underset{\sim}{b}}$ (up to a complex factor with modular one) which satisfies (2.7), and

(iii) for appropriate ordering,

$$\hat{\omega}_j \to \omega_j, \quad j = 1, \ldots, t_0, \qquad S_{\hat{t}_0} \to \sigma^2, \quad \text{as } N \to \infty.$$

In other words, $(\hat{t}_0, \{\hat{\omega}_j, j \leq \hat{t}_0\}, S_{\hat{t}_0})$ is a strongly consistent estimate of $(t_0, \{\omega_j, j \leq t_0\}, \sigma^2)$.

To prove this theorem, we need the following lemma:

LEMMA 3.1 (Petrov). Let $\{X_n, n \geq 1\}$ be a sequence of independent real random variables with zero means. Write $s_n^2 = \sum_{j=1}^{n} EX_j^2$ and $S_n = \sum_{j=1}^{n} X_j$. If

$$\liminf_{n\to\infty} s_n^2/n > 0$$

and

$$E|X_j|^{2+\delta} \leq K < \infty, \quad j \geq 1$$

for some constants K and $\delta > o$, then

$$\limsup_{n\to\infty} S_n/(2s_n^2 \log(s_n^2))^{1/2} = 1, \quad \text{a.s.}$$

For a proof, the reader is referred to Petrov (1975) and Stout (1974, p.274).

Proof of Theorem 3.1. Under the model (2.1), we have

$$\hat{\gamma}_{\ell m}^{(t)} = \frac{1}{N-t}\sum_{n=t+1}^{N} y(n-\ell)\overline{y(n-m)}$$

$$= \frac{1}{N-t}\sum_{n=t+1}^{N}\Big(\sum_{j=1}^{t_0} a_j\exp(i(n-\ell)\omega_j) + w(n-\ell)\Big)\Big(\sum_{j=1}^{t_0}\bar{a}_j\exp(-i(n-m)\omega_j) + \overline{w(n-m)}\Big)$$

$$= \sum_{j=1}^{t_0}|a_j|^2\exp(i(m-\ell)\omega_j) + \sum_{j,k=1,j\neq k}^{t_0} a_j\bar{a}_k\exp(i(m\omega_k-\ell\omega_j))\frac{1}{N-t}\sum_{n=t+1}^{N}\exp(in(\omega_j-\omega_k))$$

$$+ \sum_{j=1}^{t_0} a_j\exp(i(m-\ell)\omega_j)\frac{1}{N-t}\sum_{n=t+1}^{N}\exp(i(n-m)\omega_j)\overline{w(n-m)}$$

$$+ \sum_{j=1}^{t_0}\bar{a}_j\exp(i(m-\ell)\omega_j)\frac{1}{N-t}\sum_{n=t+1}^{N}\exp(-i(n-\ell)\omega_j)w(n-\ell)$$

$$+ \frac{1}{N-t}\sum_{n=t+1}^{N} w(n-\ell)\overline{w(n-m)}$$

$$= J_1 + J_{2N} + J_{3N} + J_{4N} + J_{5N} \quad \text{(say)}, \qquad \ell, m = 0, 1, \ldots, t. \qquad (3.2)$$

For $\omega_j \neq \omega_k$, $j \neq k$,

$$\frac{1}{N-t}\sum_{n=t+1}^{N}\exp(in(\omega_j-\omega_k)) = \frac{\exp(i(t+1)(\omega_j-\omega_k)) - \exp(i(N+1)(\omega_j-\omega_k))}{(N-t)(1-\exp(i(\omega_j-\omega_k)))} = O(\tfrac{1}{N}),$$

$$j \neq k, \quad j, k = 1, \ldots, t_0.$$

Thus

$$J_{2N} = O(\tfrac{1}{N}). \qquad (3.3)$$

By condition (2.2) and Lemma 3.1,

$$J_{3N} = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \text{ a.s.,} \qquad J_{4N} = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \text{ a.s.} \qquad (3.4)$$

as $N \to \infty$.

By the law of the iterated logarithm of M-dependence sequence,

$$J_{5N} = \begin{cases} O\left(\sqrt{\frac{\log\log N}{N}}\right), & \text{for } \ell \neq m, \\ \sigma^2 + O\left(\sqrt{\frac{\log\log N}{N}}\right), & \text{for } \ell = m, \end{cases} \quad \text{a.s.} \tag{3.5}$$

Let

$$\Gamma^{(t)} = (\gamma_{\ell m}^{(t)}), \qquad \gamma_{\ell m}^{(t)} = \sigma^2 \delta_{\ell m} + \sum_{j=1}^{t_0} |a_j|^2 \exp(i(m-\ell)\omega_j),$$

$$\ell, m = 0, 1, \ldots, t, \tag{3.6}$$

where $\delta_{\ell m}$ is the Kronecker sign.

Using (3.2)-(3.6),

$$\hat{\gamma}_m^{(t)} = \gamma_m^{(t)} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.,} \quad \ell, m = 0, 1, \ldots, t. \tag{3.7}$$

Let $\hat{\lambda}_1^{(t)} \geq \ldots \geq \hat{\lambda}_{t+1}^{(t)}$ be the eigenvalues of $\hat{\Gamma}^{(t)}$, and $\lambda_1^{(t)} \geq \ldots \geq \lambda_{t+1}^{(t)}$ be the eigenvalues of $\Gamma^{(t)}$. Then

$$\sum_{j=1}^{t+1} \lambda_j^{(t)} \hat{\lambda}_j^{(t)} \geq \operatorname{tr} \Gamma^{(t)} \hat{\Gamma}^{(t)}.$$

(see von Neumann (1937)). Hence,

$$\sum_{j=1}^{t+1} (\hat{\lambda}_j^{(t)} - \lambda_j^{(t)})^2 \leq \operatorname{tr}(\hat{\Gamma}^{(t)} - \Gamma^{(t)})^2. \tag{3.8}$$

Put

$$\underset{(t+1)\times t_0}{\Omega} = \begin{bmatrix} 1 & \cdots & 1 \\ \exp(-i\omega_1) & \cdots & \exp(-i\omega_{t_0}) \\ \cdots & \cdots & \cdots \\ \exp(-it\omega_1) & \cdots & \exp(-it\omega_{t_0}) \end{bmatrix}, \quad A = \operatorname{diag}[a_1, a_2, \ldots, a_{t_0}]. \tag{3.9}$$

Since $a_j \neq 0$, $j = 1, \ldots, t_0$, and $\omega_j \neq \omega_k$ if $j \neq k$, it is easily seen that

$$\mathrm{rank}(\Omega A) = \min(t+1, t_0), \tag{3.10}$$

$$\Gamma^{(t)} = \sigma^2 I_{t+1} + \Omega A A^* \Omega^*, \tag{3.11}$$

and Ω^* denotes the transpose of the complex conjugate of Ω. Thus

$$\begin{aligned} \lambda_{t+1}^{(t)} &> \sigma^2 \quad \text{for} \quad t < t_0, \\ \lambda_{t+1}^{(t)} &= \sigma^2 \quad \text{for} \quad t \geq t_0. \end{aligned} \tag{3.12}$$

Hence, by (3.7) and (3.8), noticing that $S_t = \hat{\lambda}_{t+1}^{(t)}$, we have

$$\lim_{N\to\infty} S_t = \lambda_{t+1}^{(t)} > \sigma^2, \quad \text{a.s.} \quad \text{for} \quad t < t_0, \tag{3.13}$$

and

$$|S_t - \sigma^2| = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \quad \text{for} \quad t \geq t_0. \tag{3.14}$$

Assume that $t < t_0$. Then by (3.13), (3.14), (2.6) and (2.5),

$$\lim_{N\to\infty}(R_{t_0} - R_t) = \sigma^2 - \lambda_{t+1}^{(t)} < 0, \quad \text{a.s.} \tag{3.15}$$

Hence, with probability one for N large

$$R_{t_0} < R_t, \quad \text{for} \quad t < t_0. \tag{3.16}$$

Now we assume that $t > t_0$. Then by (3.14), (2.6) and (2.5), with probability one for N large,

$$R_{t_0} - R_t = S_{t_0} - S_t - (t - t_0)C_N$$

$$= O\left(\sqrt{\frac{\log \log N}{N}}\right) - (t - t_0)C_N < 0. \tag{3.17}$$

Hence, with probability one for N large,

$$\hat{t}_N = t_0, \tag{3.18}$$

which establishes (i).

To prove (ii) and (iii), without loss of generality, we can assume $\hat{t}_0 = t_0$. By (3.10) and (3.11) with $t = t_0$,

$$\sigma^2 = \lambda_{t_0+1}^{(t_0)} < \lambda_{t_0}^{(t_0)} \leq \cdots \leq \lambda_1^{(t_0)}. \tag{3.19}$$

Thus, the equation

$$(\Gamma^{(t_0)} - \sigma^2 I_{t_0+1})\underset{\sim}{b} = 0, \quad \|\underset{\sim}{b}\| = 1 \tag{3.20}$$

has a unique root $\underset{\sim}{b} = (b_0, b_1, \ldots, b_{t_0})'$ (up to a complex factor with modular one). By (2.7), $\hat{\underset{\sim}{b}}$ is a unit eigenvector corresponding to the smallest eigen- $\hat{\lambda}_{t_0+1}^{(t_0)}$ of $\hat{\Gamma}^{(t_0)}$. By (3.7), (3.8) and (3.19), with probability one for N large,

$$\hat{\lambda}_{t_0+1}^{(t_0)} < \hat{\lambda}_{t_0}^{(t_0)} \leq \cdots \leq \hat{\lambda}_1^{(t_0)}, \tag{3.21}$$

which implies that the equation

$$(\hat{\Gamma}^{(t_0)} - \hat{\lambda}_{t_0+1}^{(t_0)} I_{t_0+1})\hat{\underset{\sim}{b}} = 0, \quad \|\hat{\underset{\sim}{b}}\| = 1 \tag{3.22}$$

has a unique root $\hat{\underset{\sim}{b}} = (\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_{t_0})'$ (up to a complex factor with modular one), and, with appropriate choice of this factor, we have

$$\lim_{N\to\infty} \hat{\underset{\sim}{b}} = \underset{\sim}{b}, \quad \text{a.s.} \tag{3.23}$$

Now, choose $\underset{\sim}{b} = (b_0, b_1, \ldots, b_{t_0})'$ such that $\|\underset{\sim}{b}\| = 1$ and

$$B(z) \triangleq \sum_{\ell=0}^{t_0} b_\ell z^\ell = 0 \tag{3.24}$$

has t_0 roots $\exp(i\omega_1), \exp(i\omega_2), \ldots, \exp(i\omega_{t_0})$. Then $\underset{\sim}{b}$ is the unique root of equation (3.20) (up to a complex factor with modular one) and vice-versa. Using (3.23),

$$\lim_{N\to\infty} \hat{B}(z) = B(z), \quad \text{a.s.} \tag{3.25}$$

Now, use the definition of $\hat{\rho}_j \exp(i\hat{\omega}_j)$, $j = 1, \ldots, t_0$, for appropriate ordering,

$$\lim_{N\to\infty} \hat{\rho}_j \exp(i\hat{\omega}_j) = \exp(i\omega_j), \quad \text{a.s.,} \quad \text{for} \quad j = 1, \ldots, t_0,$$

which implies

$$\lim_{N\to\infty} \hat{\omega}_j = \omega_j, \quad \text{a.s.,} \quad \text{for} \quad j = 1, \ldots, t_0. \tag{3.26}$$

Using (3.14), we establish parts (ii) and (iii) of the theorem.

Remark. The EVLP method can be easily generalized to EVFBLP along a similar line. The anoloques of the results given in this section and the next section can be proved step by step as the proof given in these two sections. Also, we can expect EVFBLP to give more accuracy when SNR is high and the sample size is small.

4. LIMITING DISTRIBUTION OF $S_{\hat{t}_0}$ AND $(\hat{\omega}_j,\ j \le \hat{t}_0)$

Since $\hat{t}_0 \to t_0$, a.s., as $N \to \infty$, we can use t_0 instead of $\hat{t}_0$ when we consider the limiting properties of various estimates involving $\hat{t}_0$. For further simplicity of notation, we will omit the subscript 0 of t_0 and simply use t for t_0. Also, we will keep all other notation defined in previous sections. In particular, the matrix Ω in (3.9) is a $(t+1) \times t$ matrix.

Throughout this section, we assume that $\{w(n)\}$ is a sequence of i.i.d. complex random variables such that

$$Ew(1) = 0, \quad E(Rw(1))^2 = E(Iw(1))^2 = \frac{1}{2}\sigma^2,$$

$$E(Rw(1)Iw(1)) = 0, \quad \text{and} \quad \mathrm{Var}(|w(1)|^2) = \alpha\sigma^4 \ \text{ with } \ \alpha > 0. \tag{4.1}$$

Here, $Rw(1)$ and $Iw(1)$ denote the real and imaginary parts of $w(1)$ respectively.

LEMMA 4.1. Suppose that condition (4.1) is satisfied. Then

$$\frac{1}{\sqrt{N-t}} \sum_{n=t+1}^{N} \exp(-i(n-\ell)\omega_j)w(n-\ell) \xrightarrow{\mathcal{D}} v_j, \quad j = 1, 2, \ldots, t, \quad \ell = 0, 1, \ldots, t,$$

$$\frac{1}{\sqrt{N-t}} \sum_{n=t+1}^{N} (|w(n-\ell)|^2 - \sigma^2) \xrightarrow{\mathcal{D}} u_0, \quad \ell = 0, 1, \ldots, t,$$

$$\frac{1}{\sqrt{N-t}} \sum_{n=t+1}^{N} w(n-\ell)\overline{w(n-m)} \xrightarrow{\mathcal{D}} u_{\ell-m}, \quad \text{if} \quad 0 \le m < \ell \le t.$$

Here v_j's and u_j's are independent of each other and

$$\begin{aligned}
&\text{(i)} && v_j \sim N_c(0, \sigma^2), && j = 1, \ldots, t,\\
&\text{(ii)} && u_0 \sim N_r(0, \alpha\sigma^4), &&\\
&\text{(iii)} && u_j \sim N_c(0, \sigma^4), && j = 1, \ldots, t.
\end{aligned} \tag{4.2}$$

Also, N_c and N_r denote complex and real normal distribution respectively.

Proof. The normality of v_j, u_0 and u_j follows directly from the central limit theorem by computing the covariance matrix and the fact that

$$\frac{1}{N}\sum_{n=1}^{N} \exp(in\omega) \to 0 \quad \text{for any } \omega \neq 0.$$

Put

$$A = \operatorname{diag}[a_1, \ldots, a_t],$$

$$\underset{(t+1)\times t}{\Omega} = \begin{pmatrix} 1 & \cdots & 1 \\ \exp(-i\omega_1) & \cdots & \exp(-i\omega_t) \\ \cdots & \cdots & \cdots \\ \exp(-it\omega_1) & \cdots & \exp(-it\omega_t) \end{pmatrix},$$

$$\underset{(t+1)\times(t+1)}{U} = \begin{pmatrix} u_0 & \bar{u}_1 & \cdots & \bar{u}_t \\ u_1 & u_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bar{u}_1 \\ u_t & \cdots & u_1 & u_0 \end{pmatrix}, \tag{4.3}$$

$$\zeta_N = \sqrt{N-t}(S_t-\sigma^2), \quad T_{Nj} = \sqrt{N-t}(\hat{\rho}_j-1), \quad \Delta_{Nj} = \sqrt{N-t}(\hat{\omega}_j-\omega_j),$$

$$\underset{\sim}{T}_n = (T_{N1}, \ldots, T_{Nt})' \quad \text{and} \quad \underset{\sim}{\Delta}_N = (\Delta_{N1}, \ldots, \Delta_{Nt})'.$$

Define $B(z) = b_t \prod_{j=1}^{t} (z - \exp(i\omega_j)) = \sum_{\ell=0}^{t} b_\ell z^\ell$ where b_t is chosen such that $\|\underset{\sim}{b}\| = 1$ for $\underset{\sim}{b} = (b_0, b_1, \ldots, b_t)'$. Write

$$D(\exp(i\omega)) = -i\frac{d}{d\omega}B(\exp(i\omega)), \quad \text{and}$$

$$G = \operatorname{diag}[D(\exp(i\omega_1)), \ldots, D(\exp(i\omega_t))]. \tag{4.4}$$

We have the following result:

THEOREM 4.1. Suppose that $\{w(n)\}$ is a i.i.d. sequence of complex random variables such that (4.1) holds. Then we have

$$\zeta_N \xrightarrow{\mathcal{D}} \zeta = \underset{\sim}{b}^* U \underset{\sim}{b}, \tag{4.5}$$

and

$$T_N + i\underset{\sim}{\Delta}_N \xrightarrow{\mathcal{D}} G^{-1}(AA^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U \underset{\sim}{b}. \tag{4.6}$$

Proof. Put $V = \mathrm{diag}[v_1, \ldots, v_t]$ and

$$\hat{H} = \sqrt{N-t}(\hat{\Gamma}^{(t)} - \Gamma^{(t)}) = (\hat{h}_{\ell m}). \tag{4.7}$$

By (3.2), (3.3) and Lemma 4.1, we have

$$\hat{H} \xrightarrow{\mathcal{D}} H = (h_{\ell m}) = \Omega(A\bar{V} + \bar{A}V)\Omega^* + U. \tag{4.8}$$

Since S_t is the smallest eigenvalue of $\hat{\Gamma}^{(t)}$, we have

$$\begin{aligned} 0 &= \det[\Gamma^{(t)} - S_t I_{t+1}] \\ &= \det[\Gamma^{(t)} - \sigma^2 I_{t+1} + \frac{1}{\sqrt{N-t}}\hat{H} - (S_t - \sigma^2)I_{t+1}] \\ &= \det[\Omega AA^*\Omega^* + \frac{1}{\sqrt{N-t}}\hat{H} - (S_t - \sigma^2)I_{t+1}]. \end{aligned} \tag{4.9}$$

Let Q be a unitary matrix such that

$$Q^*\Omega AA^*\Omega^* Q = \mathrm{diag}[\xi_1, \ldots, \xi_t, 0], \qquad \xi_1 \geq \cdots \geq \xi_t > 0.$$

Note that the last column of Q is the eigenvector of $\Gamma^{(t)}$ corresponding to the eigenvalue σ^2 of $\Gamma^{(t)}$. We can choose this column as $\underset{\sim}{b}$. Now we have

$$0 = \det\left\{\begin{pmatrix} \xi_1-(S_t-\sigma^2) & & & \\ & \ddots & & \\ & & \xi_t-(S_t-\sigma^2) & \\ & & & -(S_t-\sigma^2)\end{pmatrix} + \frac{1}{\sqrt{N-t}}\, Q^*\hat{H}Q\right\}. \qquad (4.10)$$

Since $\hat{H} \xrightarrow{\mathcal{D}} H$, by Skorohod's representation theorem (see [7]), we can assume $\hat{H} \xrightarrow{\mathcal{D}} H$, a.s. as $N \to \infty$. Multiply by $(N-t)^{1/4}$ the last row and the last column of the matrix in { } of (4.10) respectively. By (4.10), $S_t \to \sigma^2$, a.s. and noting that $\Omega^*\underset{\sim}{b} = \underset{\sim}{0}$, we have

$$\zeta_N \to \zeta = \underset{\sim}{b}^*H\underset{\sim}{b} = \underset{\sim}{b}^*U\underset{\sim}{b}, \quad \text{a.s.} \qquad (4.11)$$

Note that (4.11) reveals only that there exist some versions of ζ_N, ζ, H and U which have the same distributions as ζ_N, ζ, H and U respectively such that (4.11) holds. From this we only get (4.5). The principle of this statement also applies to the following proof of (4.6) and so on.

Since $(\hat{\Gamma}^{(t)} - S_tI_{t+1})\hat{\underset{\sim}{b}} = 0$ and $(\Gamma^{(t)} - \sigma^2I_{t+1})\underset{\sim}{b} = 0$ with the choice of $\hat{\underset{\sim}{b}}$ such that $\hat{\underset{\sim}{b}} \to \underset{\sim}{b}$, a.s., we have

$$\begin{aligned} \underset{\sim}{0} &= \sqrt{N-t}(\hat{\Gamma}^{(t)} - S_tI_{t+1})\hat{\underset{\sim}{b}} \\ &= \sqrt{N-t}\left(\Gamma^{(t)} - \sigma^2 I_{t+1} + \frac{1}{\sqrt{N-t}}\hat{H} - \frac{1}{\sqrt{N-t}}\zeta_N I_{t+1}\right)\left(\underset{\sim}{b} + \frac{1}{\sqrt{N-t}}\underset{\sim}{\eta}_N\right) \\ &= (\Gamma^{(t)} - \sigma^2 I_{t+1})\underset{\sim}{\eta}_N + (\hat{H} - \zeta_N I_{t+1})\underset{\sim}{b} + (\hat{H} - \zeta_N I_{t+1})(\hat{\underset{\sim}{b}} - \underset{\sim}{b}), \end{aligned}$$

where $\quad \underset{\sim}{\eta}_N = \sqrt{N-t}(\hat{\underset{\sim}{b}} - \underset{\sim}{b}) \triangleq (\eta_{N0}, \ldots, \eta_{Nt})'$. $\qquad$ (4.12)

Write $\underset{\sim}{\eta}_N = \underset{\sim}{\eta}_N^{(1)} + \underset{\sim}{\eta}_N^{(2)}$ such that $\underset{\sim}{\eta}_N^{(1)} \in \mu(\Omega)$ and $\underset{\sim}{\eta}_N^{(2)} \perp \mu(\Omega)$, where

$\mu(\Omega)$ denote the space spanned by all column vectors of Ω. Then $\underset{\sim}{\eta}_N^{(1)} = \Omega\underset{\sim}{\beta}_N$ for some $t \times 1$ complex vector $\underset{\sim}{\beta}_N$. Since $\hat{H} \to H$, a.s., $\zeta_N \to \zeta = \underset{\sim}{b}^* U \underset{\sim}{b}$, a.s. and $\hat{\underset{\sim}{b}} \to \underset{\sim}{b}$, a.s., we have

$$(\hat{H} - \zeta_N I_{t+1})(\hat{\underset{\sim}{b}} - \underset{\sim}{b}) \to \underset{\sim}{0}, \quad \text{a.s.} \quad \text{as } N \to \infty. \tag{4.13}$$

Note that

$$(\Gamma^{(t)} - \sigma^2 I_{t+1})\underset{\sim}{\eta}_N = \Omega AA^*\Omega^*\underset{\sim}{\eta}_N = \Omega AA^*\Omega^*\underset{\sim}{\eta}_N^{(1)} = \Omega AA^*\Omega^*\Omega\underset{\sim}{\beta}_N \tag{4.14}$$

and

$$\begin{aligned}(\hat{H} - \zeta_N I_{t+1})\underset{\sim}{b} \to (H - \zeta I_{t+1})\underset{\sim}{b} &= U\underset{\sim}{b} - (\underset{\sim}{b}^* U \underset{\sim}{b})\underset{\sim}{b} \\ &= (I_{t+1} - \underset{\sim}{b}\underset{\sim}{b}^*)U\underset{\sim}{b}, \quad \text{a.s.}\end{aligned} \tag{4.15}$$

By (4.12)-(4.15), and $I_{t+1} - \underset{\sim}{b}\underset{\sim}{b}^* = \Omega(\Omega^*\Omega)^{-1}\Omega^*$, we get

$$\underset{\sim}{\beta}_N \to -(\Omega^*\Omega)^{-1}(AA^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\underset{\sim}{b}, \quad \text{a.s.}$$

which implies that

$$\underset{\sim}{\eta}_N^{(1)} \to -\Omega(\Omega^*\Omega)^{-1}(AA^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\underset{\sim}{b}, \quad \text{a.s.} \tag{4.16}$$

Hence

$$\Omega^*\underset{\sim}{\eta}_N = \Omega^*\underset{\sim}{\eta}_N^{(1)} \to -(AA^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\underset{\sim}{b}, \quad \text{a.s.} \tag{4.17}$$

Finally, let us consider the limiting distribution of $\underset{\sim}{\Delta}_N$. Since $\hat{B}(\hat{\rho}_j \exp(i\hat{\omega}_j)) = B(\exp(i\omega_j)) = 0$, we have for $j = 1, \ldots, t$,

$$0 = \sqrt{N} \sum_{\ell=0}^{t} \hat{b}_\ell \hat{\rho}_j^\ell \exp(i\ell\hat{\omega}_j)$$

$$= \sqrt{N} \sum_{\ell=0}^{t} (b_\ell + \hat{b}_\ell - b_\ell)\{\exp(i\ell\omega_j) + (\hat{\rho}_j - 1)\exp(i\ell\hat{\omega}_j) + (\exp(i\ell\hat{\omega}_j) - \exp(i\ell\omega_j))\}$$

$$= \sum_{\ell=0}^{t} \eta_{N\ell} \exp(i\ell\omega_j) + \sum_{\ell=0}^{t} \hat{b}_\ell \{\ell\rho_{j*}^{\ell-1} T_{Nj} \exp(i\ell\hat{\omega}_j) + i\ell\Delta_{Nj} \exp(i\ell\omega_{j*})\},$$

where $\rho_{j*} \in [1, \rho_j]$ or $[\rho_j, 1]$ and $\omega_{j*} \in [\omega_j, \hat{\omega}_j]$ or $[\hat{\omega}_j, \omega_j]$. We have proved that $\hat{\underset{\sim}{b}} \to \underset{\sim}{b}$, a.s., $\hat{\rho}_j \to 1$, a.s. and $\hat{\omega}_j \to \omega_j$, a.s. for appropriate ordering and for $j = 1, \ldots, t$, so that, with probability one, we have

$$\Omega^* \underset{\sim}{\eta}_N + G(\underset{\sim}{T}_N + i\underset{\sim}{\Delta}_N) + o(\|\underset{\sim}{T}_N\| + \|\underset{\sim}{\Delta}_N\|) = \underset{\sim}{0} \qquad (4.18)$$

for large N. Since $\exp(i\omega_j)$ is a simple root of the multinomial $B(z)$, we have $D(\exp(i\omega_j)) \neq 0$ for $j = 1, 2, \ldots, t$, and G is nonsingular. From this, and (4.17)-(4.18), it follows that

$$\underset{\sim}{T}_N + i\underset{\sim}{\Delta}_N \to G^{-1}(AA^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\underset{\sim}{b}, \quad \text{a.s.} \qquad (4.19)$$

As mentioned above, this only gives the corresponding weak convergence described in (4.6). This completes the proof of Theorem 4.1.

5. FURTHER DISCUSSION ON MODIFIED FBLP METHOD

Consider the model (2.1) and suppose that (4.1) holds. In this section we discuss the problem of estimating the frequencies when t_0 is known. It is emphasized that some of the notations of this section may be different from those of the previous sections.

Nutall (1976) and Ulrych and Clayton (1976) developed the method of forward-backward linear prediction (FBLP), which works well when the signal-to-noise ratio (SNR) is sufficiently high. However, it becomes false in the case of low SNR. Tufts and Kumaresan (1982) proposed a modified FBLP (MFBLP) method. They showed by simulation, that the estimation of the frequencies has been greatly improved by MFBLP when the SNR is relatively low. In this section we shall give a theoretical analysis to show why the MFBLP improves when SNR is low and the sample size is large. Also, we shall suggest a further modification to this method.

First, let us describe the FBLP procedure. The reader should note that it is assumed that the true number of signals is known in this procedure.

The linear prediction coefficient vector $\hat{\underset{\sim}{b}}^{(t)}(t \geq t_0)$ is defined as the one which minimizes

$$\sum_{n=t+1}^{N} |y(n) + \sum_{\ell=1}^{t} \hat{b}_\ell^{(t)} y(n-\ell)|^2 + \sum_{n=1}^{N-t} |y^*(n) + \sum_{\ell=1}^{t} \hat{b}_\ell^{(t)} y^*(n+\ell)|^2, \qquad (5.1)$$

where $\hat{b}_\ell^{(t)}$, $\ell = 1,2,\ldots,t$, are the components of $\hat{\underset{\sim}{b}}^{(t)}$.

In case the solution is not unique, we select the one which also minimizes

$$\sum_{\ell=1}^{t} |\hat{b}_\ell^{(t)}|^2. \qquad (5.2)$$

In light of this, we call the solution $\hat{\underset{\sim}{b}}^{(t)}$ least normed prediction (LNP) coefficient vector.

We can construct the transfer function of the prediction error filter

$$\hat{B}^{(t)}(z) = 1 + \sum_{\ell=1}^{t} \hat{b}_{\ell}^{(t)} z^{-\ell}. \tag{5.3}$$

Let $\hat{\rho}_{\ell}^{(t)}\exp(i\hat{\omega}_{\ell}^{(t)})$, $\hat{\rho}_{\ell}^{(t)} > 0$, $\hat{\omega}_{\ell}^{(t)} \in [0,2\pi)$, $\ell = 1, 2, \ldots, t_0$ be the first t_0 roots of $\hat{B}^{(t)}(z)$ which are closest to the unit circle in the complex plane among the t roots, and take $\hat{\omega}_{\ell}^{(t)}$, $\ell = 1, 2, .., t_0$ as the estimates of ω_{ℓ}'s.

To introduce the MFCLP method, write

$$\hat{Y} = \begin{pmatrix} y(t), & y(t-1), & \ldots, & y(1) \\ y(t+1), & y(t), & \ldots, & y(2) \\ \ldots, & \ldots, & \ldots, & \ldots \\ \ldots, & \ldots, & \ldots, & \ldots \\ \ldots, & \ldots, & \ldots, & \ldots \\ y(N-1), & y(N-2), & \ldots, & y(N-t) \\ y^*(2), & y^*(3), & \ldots, & y^*(t+1) \\ y^*(3), & y^*(4), & \ldots, & y^*(t+2) \\ \ldots, & \ldots, & \ldots, & \ldots, \\ \ldots, & \ldots, & \ldots, & \ldots, \\ \ldots, & \ldots, & \ldots, & \ldots, \\ y^*(N-t+1), & y^*(N-t+2), & \ldots, & y^*(N) \end{pmatrix}$$

$$\hat{\underset{\sim}{b}}^{(t)} = \begin{pmatrix} \hat{b}_1^{(t)} \\ \cdot \\ \cdot \\ \cdot \\ \hat{b}_t^{(t)} \end{pmatrix}$$

and

$$\hat{\underset{\sim}{h}} = (y(t+1),\ y(t+2),\ \ldots,\ y(N),\ \ y^*(1),\ \ldots,\ y^*(N-t))'.$$

It is well-known that the LNP is

$$\hat{\underset{\sim}{h}}^{(t)} = -(\hat{Y}^*\hat{Y})^+\hat{Y}^*\hat{\underset{\sim}{h}} \tag{5.4}$$

where $(\hat{Y}^*\hat{Y})^+$ denotes the Moore-Penrose inverse of $(\hat{Y}^*\hat{Y})$.

Now, let $(2N-2t)^{-1}\hat{Y}^*\hat{Y}$ have the following spectral decomposition

$$(2N-2t)^{-1}\hat{Y}^*\hat{Y} = \sum_{j=1}^{t} \hat{\lambda}_j \hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}_j^*. \tag{5.5}$$

As shown in the sequel, with probability one for large N, the above matrix is positive definite. Hence, the FBLP gives the LNP as follows:

$$\hat{\underset{\sim}{b}}^{(t)} = \sum_{j=1}^{t} \hat{\lambda}_j^{-1} (\hat{\underset{\sim}{u}}_j^* \hat{\underset{\sim}{\gamma}}) \hat{\underset{\sim}{u}}_j, \tag{5.6}$$

where

$$\hat{\underset{\sim}{\gamma}} = -(2N-2t)^{-1}\hat{Y}^*\hat{\underset{\sim}{h}}. \tag{5.7}$$

If we use

$$\hat{\underset{\sim}{b}}_* = \sum_{j=1}^{t_0} \hat{\lambda}_j^{-1} (\hat{\underset{\sim}{u}}_j^* \hat{\underset{\sim}{\gamma}}) \hat{\underset{\sim}{u}}_j \tag{5.8}$$

instead of $\hat{\underset{\sim}{b}}^{(t)}$ in the above FBLP procedure, we get the MFBLP method.

In this section, we propose to use

$$\hat{\underset{\sim}{b}} = \sum_{j=1}^{t_0} (\hat{\lambda}_j - \hat{\lambda}_t)^{-1} (\hat{\underset{\sim}{u}}_j^* \hat{\underset{\sim}{\gamma}}) \hat{\underset{\sim}{u}}_j, \quad t > t_0, \tag{5.9}$$

instead of $\hat{\underset{\sim}{b}}_*$ in the MFBLP procedure. We will establish the strong consistency of the estimates of the frequencies for this method under the condition (4.1).

Now, we consider the case without noise. To distinguish this case from the case with noise, we drop out the superscript "$\wedge$" in all notations.

Write

$$Q = \begin{pmatrix} a_1\exp(it\omega_1), & \ldots, & a_{t_0}\exp(it\omega_{t_0}) \\ a_1\exp(i(t+1)\omega_1), & \ldots, & a_{t_0}\exp(i(t+1)\omega_{t_0}) \\ \ldots, & \ldots, & \ldots \\ a_1\exp(i(N-1)\omega_1), & \ldots, & a_{t_0}\exp(i(N-1)\omega_{t_0}) \\ \bar{a}_1\exp(-2i\omega_1), & \ldots, & \bar{a}_{t_0}\exp(-2i\omega_{t_0}) \\ \bar{a}_1\exp(-3i\omega_1), & \ldots, & \bar{a}_{t_0}\exp(-3i\omega_{t_0}) \\ \ldots, & \ldots, & \ldots, \\ \bar{a}_1\exp(-i(N-t+1)\omega_1), & \ldots, & \bar{a}_{t_0}\exp(-i(N-t+1)\omega_{t_0}) \end{pmatrix}$$

and

$$\Omega = \begin{pmatrix} 1 & \exp(-i\omega_1) & \ldots & \exp(-i(t-1)\omega_1) \\ 1 & \exp(-i\omega_2) & \ldots & \exp(-i(t-1)\omega_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \exp(-i\omega_{t_0}) & \ldots & \exp(-i(t-1)\omega_{t_0}) \end{pmatrix}.$$

Then

$$Y = Q\Omega \tag{5.10}$$

and

$$\underset{\sim}{h} = Q\underset{\sim}{\Omega}_0, \tag{5.11}$$

where
$$\underset{\sim}{\Omega}_0 = (\exp(i\omega_1), \ldots, \exp(i\omega_{t_0}))'.$$

Substituting (5.10), (5.11) into (5.4), we get

$$\underset{\sim}{b}^{(t)} = -(\Omega^* Q^* Q\Omega)^+ \Omega^* Q^* Q\underset{\sim}{\Omega}_0. \tag{5.12}$$

Note that Q and Ω have rank t_0. By simple computation we find

$$\underset{\sim}{b}^{(t)} = -\Omega^*(\Omega\Omega^*)^{-1}\underset{\sim}{\Omega}_0. \tag{5.13}$$

Let $B^{(t)}(z)$ be the function constructed according (5.3) with the vector $\underset{\sim}{b}^{(t)}$. It is easy to verify that $B^{(t)}(\exp(i\omega_\ell)) = 0$, $\ell = 1, 2, \ldots, t_0$. It is interesting that the LNP, in the noiseless case, is independent of the amplitudes.

Next, we shall consider the case when the noise arises. Let

$$W = \begin{bmatrix} w(t), & w(t-1), & \ldots, & w(1) \\ w(t+1), & w(t), & \ldots, & w(2) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ w(N-1), & w(N-2), & \ldots, & w(N-t) \\ w^*(2), & w^*(3), & \ldots, & w^*(t+1) \\ w^*(3), & w^*(4), & \ldots, & w^*(t+2) \\ \cdots & \cdots & \cdots & \cdots \\ w^*(N-t+1), & w^*(N-t+2), & \ldots, & w^*(N) \end{bmatrix}$$

and

$$\underset{\sim}{W}_0 = [w(t+1), w(t+2), \ldots, w(N), w^*(1), w^*(2), \ldots, w^*(N-t)]'.$$

Then we have

$$\hat{Y} = Q\Omega + W \tag{5.14}$$

and

$$\hat{\underset{\sim}{h}} = Q\underset{\sim}{\Omega}_0 + \underset{\sim}{W}_0. \tag{5.15}$$

By (5.14) we have

$$\hat{Y}^*\hat{Y} = \Omega^*Q^*Q\Omega + W^*Q\Omega + \Omega^*Q^*W + W^*W. \tag{5.16}$$

Under the conditions of Theorem 4.1, we can similarly prove that

$$(2N-2t)^{-1}\{W^*Q\Omega + \Omega^*Q^*W\} = O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.17}$$

and

$$(2N-2t)^{-1}W^*W - \sigma^2 I_t = O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.18}$$

Hence

$$(2N-2t)^{-1}\hat{Y}^*\hat{Y} = (2N-2t)^{-1}\Omega^*Q^*Q\Omega + \sigma^2 I_t + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.19}$$

Similarly we can prove

$$(2N-2t)^{-1}\hat{Y}^*\underset{\sim}{\hat{h}} = (2N-2t)^{-1}\Omega^*Q^*Q\underset{\sim}{\Omega}_0 + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.20}$$

Put

$$A = \mathrm{diag}[|a_1|^2, \ldots, |a_{t_0}|^2].$$

It is easily seen that

$$(2N-2t)^{-1}Q^*Q = A + O\left(\frac{1}{N}\right), \tag{5.21}$$

and

$$(2N-2t)^{-1}\Omega^*Q^*Q\Omega = \Omega^*A\Omega + O\left(\frac{1}{N}\right),$$

$$(2N-2t)^{-1}\Omega^*Q^*Q\underset{\sim}{\Omega}_0 = \Omega^*A\underset{\sim}{\Omega}_0 + O\left(\frac{1}{N}\right). \tag{5.22}$$

Therefore,

$$(2N-2t)^{-1}\hat{Y}^*\hat{Y} = \Omega^*A\Omega + \sigma^2 I_t + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.23}$$

and

$$(2N-2t)^{-1}\hat{Y}^*\hat{h} = \Omega^* A \Omega_0 + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s.} \tag{5.24}$$

Note that $\underset{\sim}{b}^{(t)}$ is independent of N in view of (5.13). Hence, by (5.12) and (5.22) we get

$$\begin{aligned}\underset{\sim}{b}^{(t)} &= -(\Omega^* A\Omega)^+ \Omega^* A \underset{\sim}{\Omega}_0 \\ &= (\Omega^* A\Omega)^+ \underset{\sim}{\beta}.\end{aligned} \tag{5.25}$$

Here we have

$$\underset{\sim}{\beta} = -\Omega^* A \underset{\sim}{\Omega}_0. \tag{5.26}$$

Note that the rank of the $t \times t$ matrix $\Omega^* A\Omega$ is t_0. So the matrix $\Omega^* A\Omega$ has the following spectral decomposition

$$\Omega^* A\Omega = \sum_{j=1}^{t_0} \zeta_j \underset{\sim}{v}_j \underset{\sim}{v}_j^*, \tag{5.27}$$

where $\zeta_1 \geq \zeta_2 \geq \cdots \geq \zeta_{t_0}$ are the non-zero eigenvalues of $\Omega^* A\Omega$, and $\underset{\sim}{v}_j$'s are the corresponding eigenvectors.

Applying Lemma 2.1 in Bai-Krishnaiah-Zhao (1985), by (5.23), we obtain

$$\begin{aligned}\hat{\lambda}_j &= \zeta_j + \sigma^2 + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s., for } j \leq t_0, \\ \hat{\lambda}_j &= \sigma^2 + O\left(\sqrt{\frac{1}{N}\log\log N}\right), \quad \text{a.s., for } j = t_0+1, \ldots, t.\end{aligned} \tag{5.28}$$

By (5.5), (5.23), (5.27) and (5.28), we get

$$\begin{aligned}\sum_{j=1}^{t} (\lambda_j - \sigma^2)\hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}_j^* &= \sum_{j=1}^{t_0} \zeta_j \underset{\sim}{v}_j \underset{\sim}{v}_j^* + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \\ \sum_{j=1}^{t_0} \zeta_j \hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}_j^* &= \sum_{j=1}^{t_0} \zeta_j \underset{\sim}{v}_j \underset{\sim}{v}_j^* + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.}\end{aligned} \tag{5.29}$$

and

$$\sum_{j=1}^{t_0} \hat{\lambda}_j \hat{u}_j \hat{u}_j^* + \sigma^2 \sum_{j=t_0+1}^{t_0} \hat{u}_j \hat{u}_j^* = \sum_{j=1}^{t_0} (\zeta_j + \sigma^2) v_j v_j^*$$

$$+ \sigma^2 \sum_{j=t_0+1}^{t} v_j v_j^* + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.30}$$

Take v_j, $j = t_0+1, \ldots, t$ as the rest unit eigenvectors of $\Omega^* A \Omega$ such that $v_j^* v_\ell = 0$ for $j \neq \ell$, $j, \ell = 1, \ldots, t$. Put

$$\underset{t \times t_0}{\hat{U}_1} = (\hat{u}_1, \ldots, \hat{u}_{t_0}), \qquad \underset{t \times (t-t_0)}{U_2} = (\hat{u}_{t_0+1}, \ldots, \hat{u}_t),$$

$$\underset{t \times t_0}{V_1} = (v_1, \ldots, v_{t_0}), \qquad \underset{t \times (t-t_0)}{V_2} = (v_{t_0+1}, \ldots, v_t),$$

and

$$\underset{t \times (t-t_0)}{\hat{U}_2} = \underset{t \times t_0}{V_1} G_{1N} + \underset{t \times (t-t_0)}{V_2} G_{2N},$$

where G_{1N} and G_{2N} are $t_0 \times (t-t_0)$ and $(t-t_0) \times (t-t_0)$ matrices respectively. By (5.29),

$$\sum_{j=1}^{t_0} \zeta_j \hat{u}_j (\tilde{u}_j^* v_k) = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \quad \text{for } k = t_0+1, \ldots, t.$$

Thus

$$\hat{U}_1^* V_2 = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.31}$$

By (5.30) and (5.31),

$$\sigma^2 V_2^* \hat{U}_2 \hat{U}_2^* V_2 = \sigma^2 I_{t-t_0} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.32}$$

Thus

$$G_{2N}G^*_{2N} = I_{t-t_0} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.}$$
$$G^*_{2N}G_{2N} = I_{t-t_0} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.33}$$

Similar to (5.31), we have

$$V^*_1\hat{U}_2 = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.34}$$

which implies that

$$G_{1N} = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.35}$$

From the above argument, it follows that

$$\sum_{j=t_0+1}^{t} \hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}^*_j = \hat{U}_2\hat{U}^*_2 = V_2V^*_2 + O\left(\sqrt{\frac{\log\log N}{N}}\right)$$
$$= \sum_{j=t_0+1}^{t} \underset{\sim}{v}_j \underset{\sim}{v}^*_j + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.36}$$

$$\sum_{j=t_0+1}^{t} \sigma^{-2}\hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}^*_j = \sigma^{-2}\sum_{j=t_0+1}^{t} \underset{\sim}{v}_j \underset{\sim}{v}^*_j + O\left(\sigma^{-2}\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.37}$$

and

$$\sum_{j=t_0+1}^{t} \hat{\lambda}_j^{-1}\hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}^*_j = \sigma^{-2}\sum_{j=t_0+1}^{t} \underset{\sim}{v}_j \underset{\sim}{v}^*_j + O\left(\sigma^{-2}\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.38}$$

Here we write the factor σ^{-2} in the remainder in order to compare the above methods in the sequel. Using similar argument, we can prove that

$$\sum_{j=1}^{t_0} \hat{\lambda}_j^{-1}\hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}^*_j = \sum_{j=1}^{t_0} (\zeta_j + \sigma^2)^{-1} \underset{\sim}{v}_j \underset{\sim}{v}^*_j + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.39}$$

and

$$\sum_{j=1}^{t_0} (\hat{\lambda}_j - \hat{\lambda}_t)^{-1}\hat{\underset{\sim}{u}}_j \hat{\underset{\sim}{u}}^*_j = \sum_{j=1}^{t_0} \zeta_j^{-1} \underset{\sim}{v}_j \underset{\sim}{v}^*_j + O\left(\sqrt{\frac{\log\log N}{N}}\right), \text{ a.s.} \tag{5.40}$$

By (5.7), (5.24) and (5.26),

$$\hat{\gamma} = \underset{\sim}{\beta} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.41}$$

Note that by (5.25) and (5.27), we see that

$$\sum_{j=1}^{t_0} \zeta_j^{-1} (\underset{\sim}{v}_j^* \underset{\sim}{\beta}) \underset{\sim}{v}_j = (\Omega^* A \Omega)^+ \underset{\sim}{\beta} = \underset{\sim}{b}^{(t)}. \tag{5.42}$$

By (5.27), we have

$$V_2^* \Omega^* A \Omega = 0, \quad \text{and} \quad V_2^* \Omega^* A = 0. \tag{5.43}$$

From this and $\underset{\sim}{\beta} = -\Omega^* A \underset{\sim}{\Omega}_0$, it follows that

$$\sum_{j=t_0+1}^{t} (\underset{\sim}{v}_j^* \underset{\sim}{\beta}) \underset{\sim}{v}_j = \underset{\sim}{0}. \tag{5.44}$$

Thus, by (5.6), (5.8), (5.9), (5.38)-(5.42) and (5.44), we have

$$\hat{\underset{\sim}{b}}^{(t)} = \sum_{j=1}^{t_0} (\zeta_j + \sigma^2)^{-1} (\underset{\sim}{v}_j^* \underset{\sim}{\beta}) \underset{\sim}{v}_j + O\left(\sigma^{-2} \sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.45}$$

$$\hat{\underset{\sim}{b}}_* = \sum_{j=1}^{t_0} (\zeta_j + \sigma^2)^{-1} (\underset{\sim}{v}_j^* \underset{\sim}{\beta}) \underset{\sim}{v}_j + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.46}$$

and

$$\hat{\underset{\sim}{b}} = \underset{\sim}{b}^{(t)} + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.47}$$

Put $\hat{\underset{\sim}{b}} = (\hat{b}_1, \ldots, \hat{b}_t)'$, $\hat{B}(z) = 1 + \sum_{\ell=1}^{t} \hat{b}_\ell z^{-\ell}$ and let $\hat{\rho}_j \exp(i\hat{\omega}_j)$, $j = 1, 2, \ldots, t$, be the roots of $\hat{B}(z)$, where $\hat{\rho}_1 \geq \hat{\rho}_2 \geq \cdots \geq \hat{\rho}_t > 0$. As pointed out earlier, ω_j, $j = 1, \ldots, t_0$, are the roots of $B^{(t)}(z)$. Kumaresan (1982) has shown that the particular choice of $\underset{\sim}{b}^{(t)}$ places the $t - t_0$ extraneous zeros of the prediction-error filter transfer function $B^{(t)}(z)$ inside of the unit circle. Thus, by (5.47) we obtain that for appropriate ordering of $\hat{\omega}_1, \ldots, \hat{\omega}_{t_0}$,

$$\hat{\omega}_j = \omega_j + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \quad \text{for } j = 1, \ldots, t_0, \tag{5.48}$$

which implies the strong consistency of the third method by us.

By (5.45) and (5.46), we see that the FBLP procedure and the MFBLP procedure have the same asymptotic behavior.

Denote by $\tilde{B}(z)$ the transfer function constructed according to (5.3) with $\hat{\underset{\sim}{b}}^{(t)}$ replaced by $\sum_{j=1}^{t_0} (\zeta_j + \sigma^2)^{-1} (\underset{\sim}{v}_j^* \underset{\sim}{\beta}) \underset{\sim}{v}_j$. In the same way, we can construct $\hat{B}_*(z)$ using $\hat{\underset{\sim}{b}}_*$ instead of $\hat{\underset{\sim}{b}}^{(t)}$ in (5.3). By (5.46) we know that

$$\hat{B}_*(\exp(i\omega_j)) = \tilde{B}(\exp(i\omega_j)) + O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.49}$$

for $j = 1, \ldots, t_0$. In general, $\tilde{B}(\exp(i\omega_j)) \neq B^{(t)}(\exp(i\omega_j)) = 0$ for $j = 1, \ldots, t_0$ in view of (5.42), so that MFBLP method is not consistent. But we know that $\tilde{B}(z)$ reduces to $B^{(t)}(z)$ when $\sigma^2 = 0$. Thus when σ^2 is small enough (with respect to λ_j, $j = 1, \ldots, t_0$), $\tilde{B}(\exp(i\omega_j)) \approx 0$ and $\hat{B}_*(\exp(i\omega_j))$ ≈ 0 for $j = 1, 2, \ldots, t_0$.

In the same way,

$$\hat{B}^{(t)}(\exp(i\omega_j)) = \tilde{B}(\exp(i\omega_j)) + O\left(\sigma^{-2}\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{5.50}$$

for $j = 1, \ldots, t_0$. When σ^2 is small, the main term of the above expression $\tilde{B}(\exp(i\omega_j)) \approx 0$ for $j \le t_0$. This means that when σ^2 is small and fixed, the FBLP procedure can estimate the true frequencies well in the large sample case. However, when N is not large, the remainder in (5.50) would bring much random effect on the zeros of the prediction-error filter transfer function $\hat{B}^{(t)}(z)$. As shown by the simulation results given in Fig. 4 and Fig. 5 in Tufts and Kumaresan's paper (1982), the main random effect is imposed on the $t - t_0$ extraneous zeros of $\hat{B}^{(t)}(z)$. If we drop $\sum_{j=t_0+1}^{t} \hat{\lambda}_j^{-1}(\hat{\underset{\sim}{u}}_j^* \hat{\underset{\sim}{\gamma}})\hat{\underset{\sim}{u}}_j$ when we construct the transfer function, as done by Tufts and Kumaresan, the estimates of the true frequencies certainly can be improved. This is why the MFSLP procedure is better than the FBLP procedure, as shown by the above simulation results.

Thus, we establish the following.

THEOREM 5.1. Suppose that (4.1) holds in the model (2.1). For the third procedure proposed by us in this section, and for appropriate ordering of $\hat{\omega}_1, \ldots, \hat{\omega}_{t_0}$, we have

$$\hat{\omega}_j = \omega_j + O\left(\sqrt{\frac{\log \log N}{N}}\right), \quad \text{a.s.}$$

for $j = 1, \ldots, t_0$.

REFERENCES

[1] Bai, Z.D., Krishnaiah, P.R. and Zhao, L.C. (1985). On rates on convergence of fast detection criteria in signal processing with white noise. Tech. Rept. 85-45, Center for Multivariate Analysis, University of Pittsburgh.

[2] Bai, Z.D. and Liang, W.Q. (1985). Strong representation of weak convergence. Tech. Rept. 85-29, Center for Multivariate Analysis, University of Pittsburgh.

[3] Kumaresan, R. (1982). Estimating the parameters of exponentially damped/undamped sinusoidal signals in noise. Ph.D. dissertation, University of Rhode Island, Kingston, August 1982.

[4] Liggett, W.S. (1933). Passive sonar: fitting models to multiple time series. In *Signal Processing*, J.W.R. Griffiths et al, Eds., New York: Academic Press.

[5] Nuttall, A.H. (1976). Spectral analysis of a univariate process with bad data points via maximum entropy and linear prediction techniques. In Naval Underwater System Center (NUSC) Scientific and Engineering Studies, Spectral Estimation, NUSC, New London, CT.

[6]. Petrov, V.V. (1975). *Sums of Independent Random Variables*. Springer-Verlag.

[7] Rao, C.R. (1986). Some recent results in signal detection. Tech. Rept. 86-23, Center for Multivariate Analysis, University of Pittsburgh.

[8] Skorokhod, A.V. (1956). Limit theorems for stochastic process. *Theory of Prob. and its Appl.*, Vol. 1, No. 3, 261-290.

[9] Stout, W.F. (1974). *Almost Sure Convergence*. Academic Press, New York.

[10] Tufts, D.W. and Kumaresan, R. (1980). Improved spectral resolution. *Proc. IEEE* (Lett.), Vol. 68, No. 3, 419-420.

[11] Tufts, D.W. and Kumaresan, R. (1980). Improved spectral resolution II, in *Proc. IEEE*, 592-597.

[12] Tufts, D.W. and Kumaresan, R. (1982). Estimation of frequencies of multiple Sinusoids: Making linear prediction perform like maximum likelihood, in *Proc. IEEE*, Vol. 70, No. 9, 975-989.

[13] Ulrych, T.J. and Clayton, R.W. (1976). Time series modeling and maximum entropy. *Phys. Earth Planet Interiors*, Vol. 12, 188-200.

[14] von Neumann, J. (1937). Some matrix-inequalities and metrization of metric space. *Tomsk University*, Rev. 1, 286-300.

END

12-87

DTIC